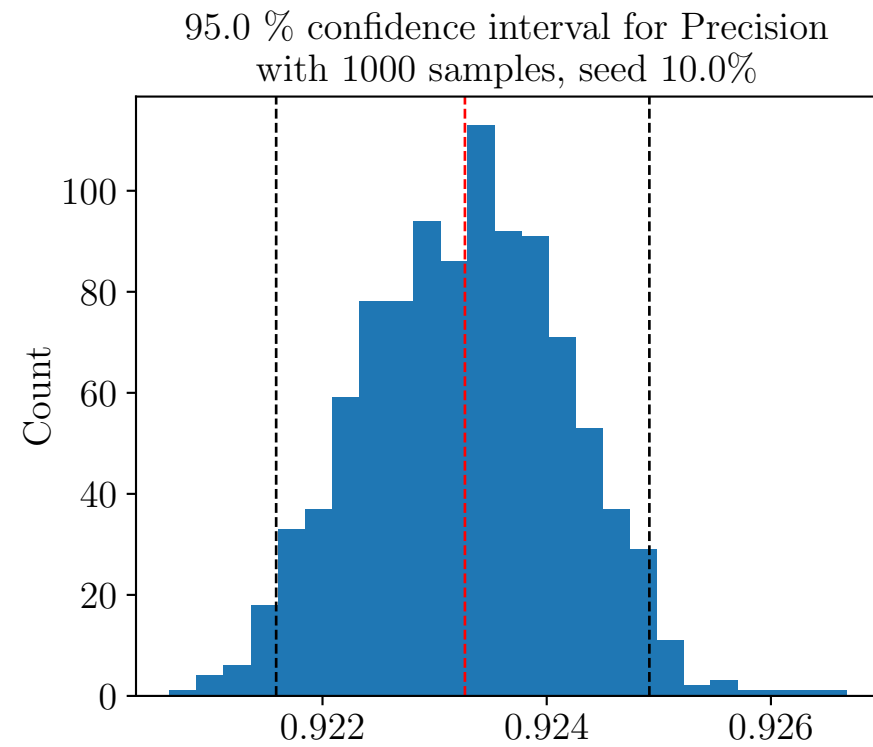
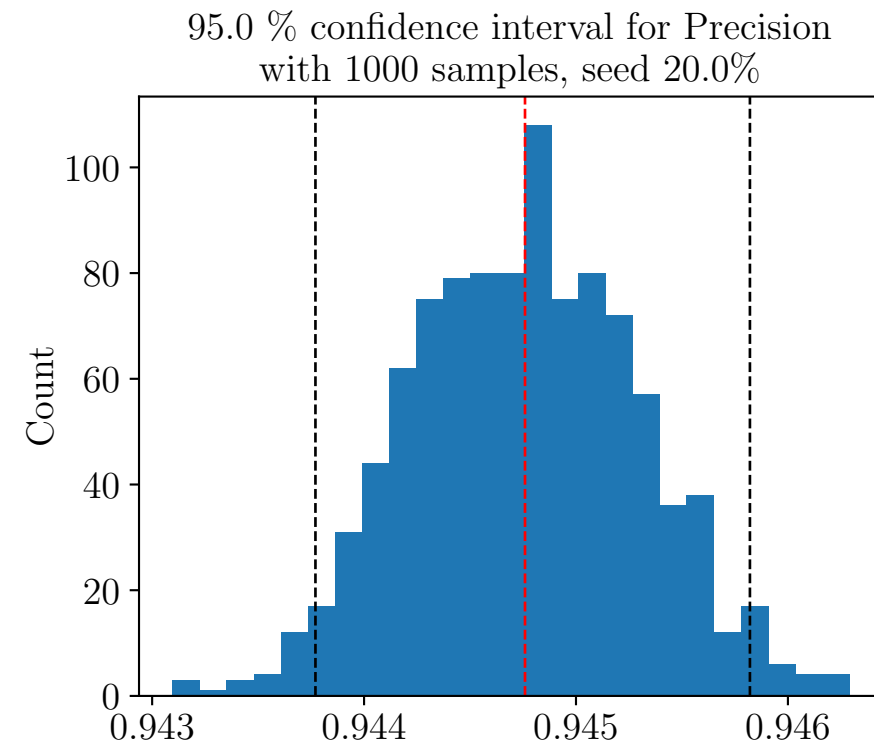


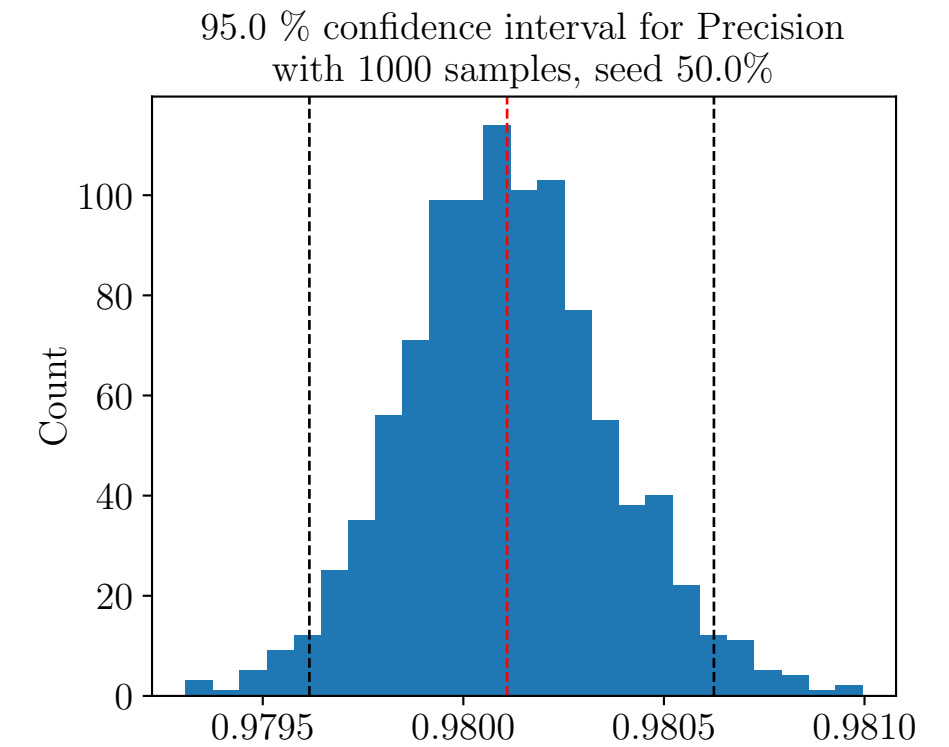
Computed confidence intervals at the last iteration, for different seeds/metrics



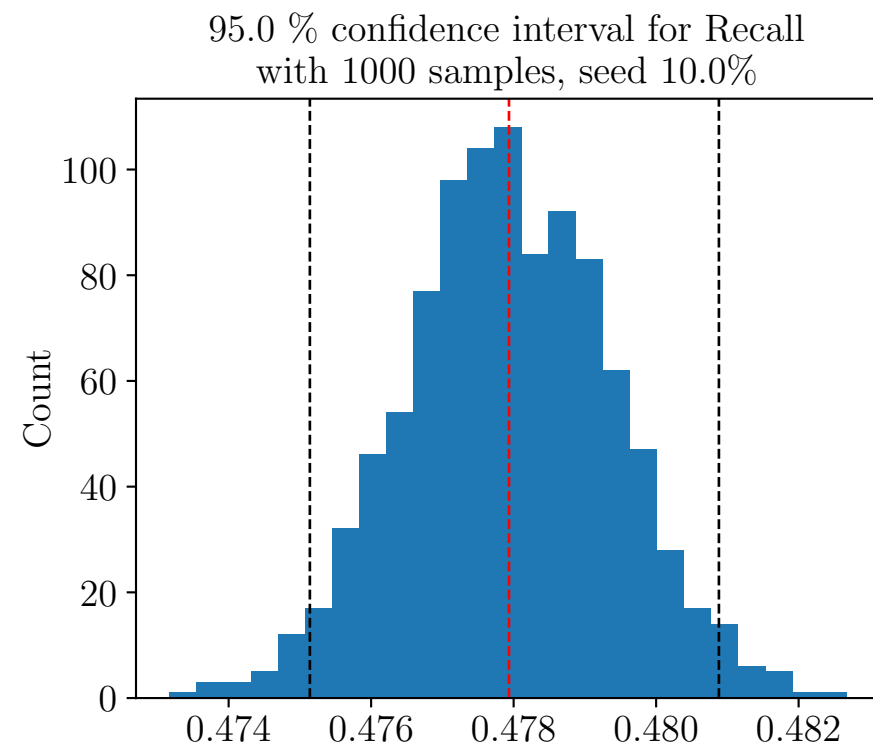
Computed bootstrap means for precision with seed 10.0%



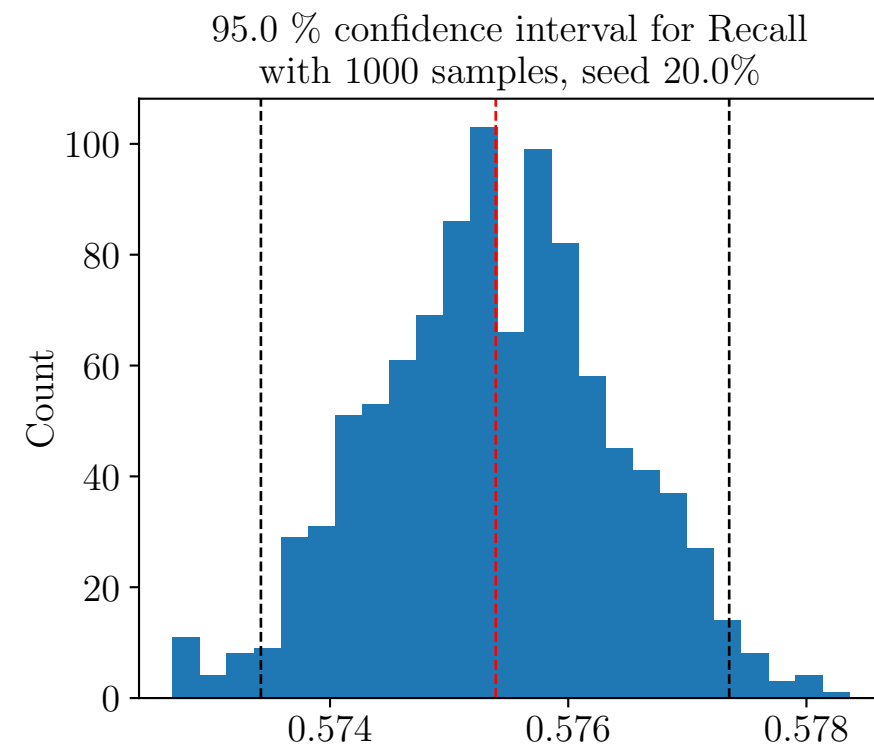
Computed bootstrap means for precision with seed 20.0%



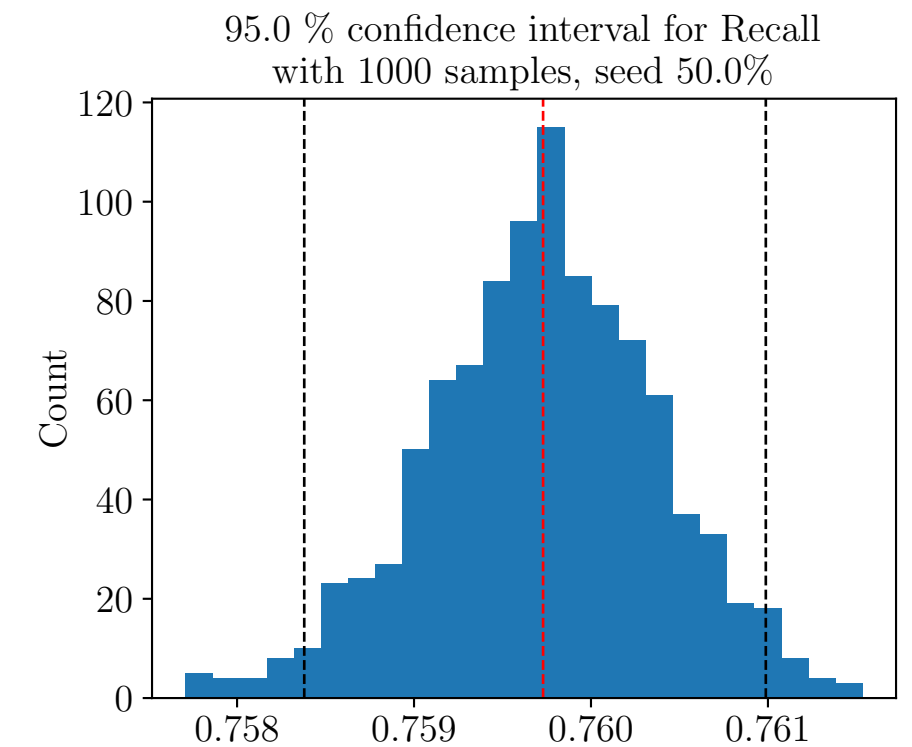
Computed bootstrap means for precision with seed 50.0%



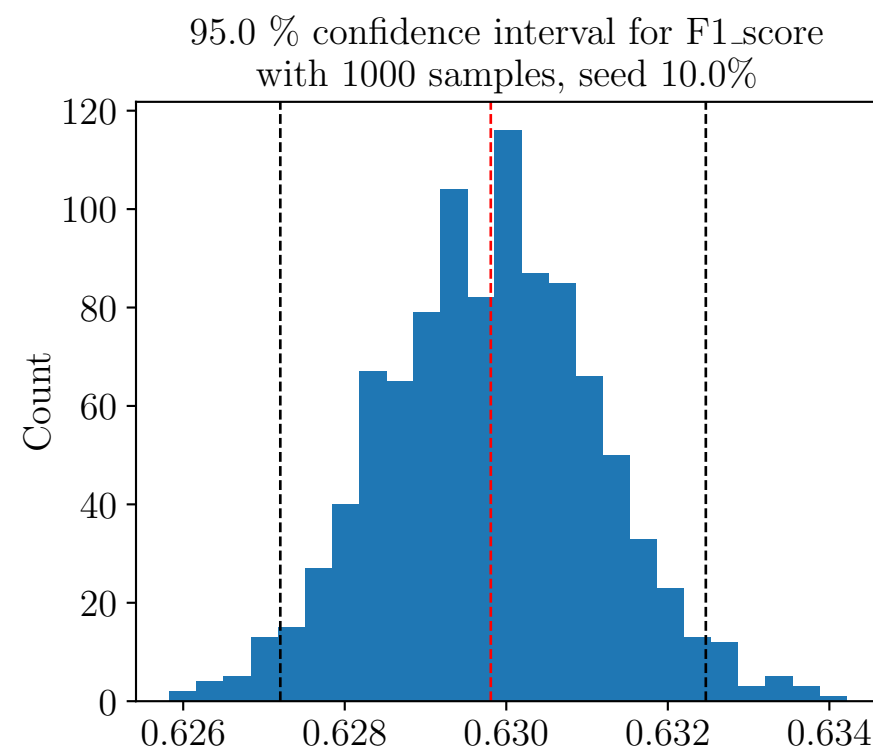
Computed bootstrap means for recall with seed 10.0%



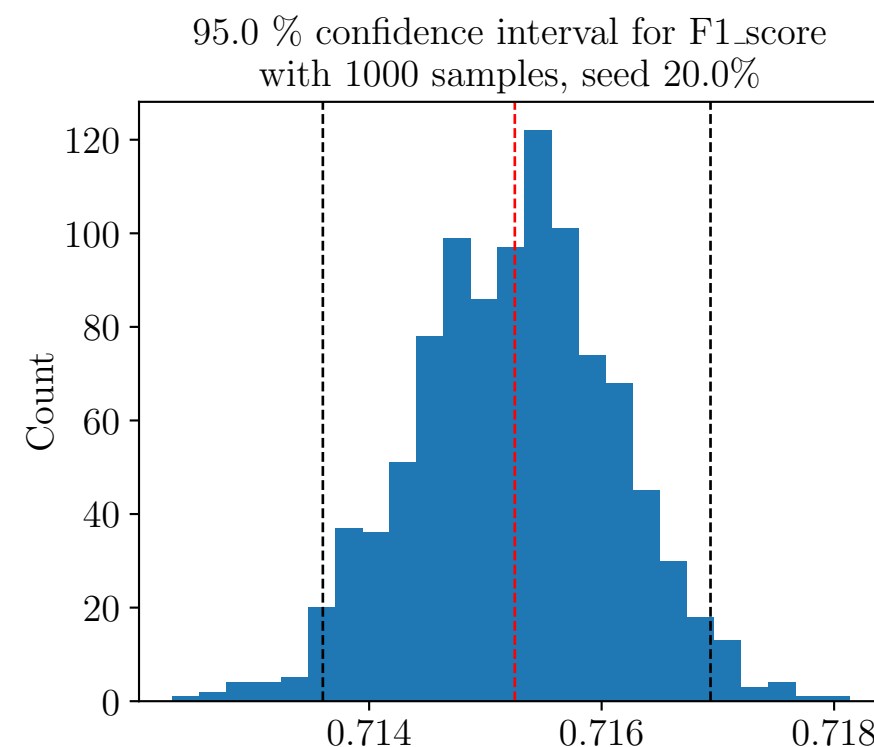
Computed bootstrap means for recall with seed 20.0%



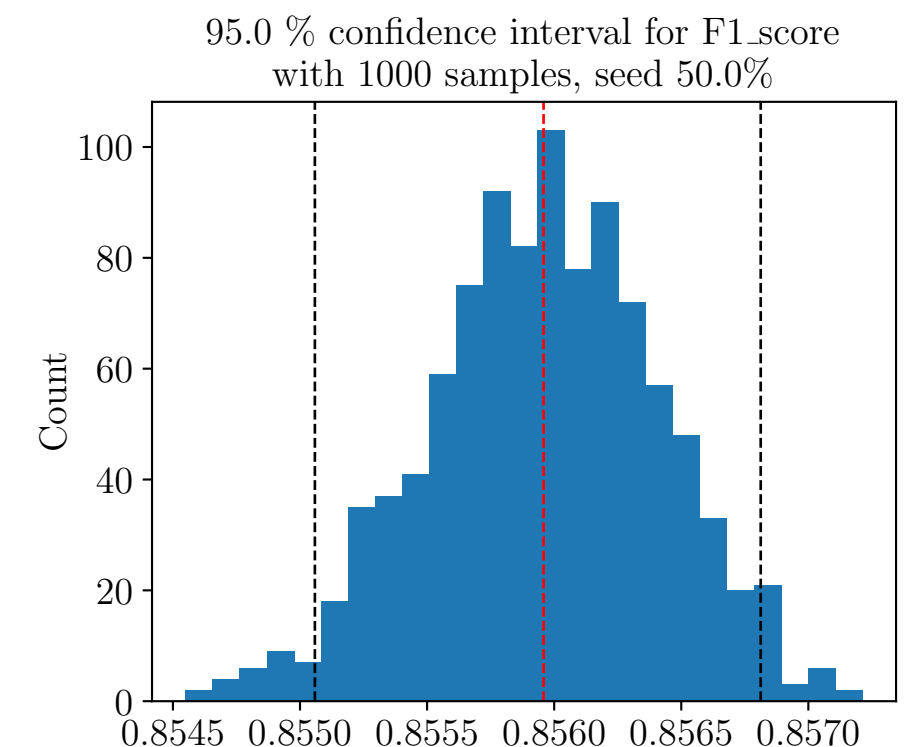
Computed bootstrap means for recall with seed 50.0%



Computed bootstrap means for f1_score with seed 10.0%



Computed bootstrap means for f1_score with seed 20.0%



Computed bootstrap means for f1_score with seed 50.0%